# Refinement of the gonadotropin releasing hormone receptor I homology model by applying molecular dynamics

Haralambos Tzoupis[1], Agathi Nteli[1], Jamie Platts[2], Efi Mantzourani[3] and Theodore Tselios[1,*]

[1]*Laboratory of Organic Chemistry, Department of Chemistry, University of Patras, Rio, Greece.*

[2]*School of Chemistry, Cardiff University, CF10 3AT, Wales, UK.*

[3]*Cardiff School of Pharmacy and Pharmaceutical Sciences, Cardiff University, CF10 3NB, Wales, UK.*

**Abstract**

Sexual maturation of human cells in ovaries and prostate is linked to the biochemical cascade initiated by the activation of cell receptors through the binding of Gonadotropin Releasing Hormone (GnRH). The GnRH receptors (GnRHR) are part of the rhodopsin G-protein-coupled receptor (GPCR) family and consist of seven trans–membrane helical domains connected via extra– and intra–cellular segments. The GnRH–GnRHR complex has been implicated in various forms of prostate and ovarian cancer. The lack of any structural data about the GnRH receptor impedes the design of antagonists for use in cancer treatment. The aim of the study is to devise a model of GnRHR to be used further for the design of improved peptide / non-peptide GnRH analogues and, to our knowledge provide new structural information regarding the extracellular loop 2 (ECL2) that acts a regulator of ligand entry to GnRHR. The common structural characteristics, of the members of the rhodopsin family of GPCRs, have been employed for the construction of a homology model for GnRHR. Structural information from the human β2–adrenergic receptor, as well as rhodopsins have been used in order to create a theoretical model for GnRHR. Furthermore, molecular dynamics (MD) simulations have been employed for the refinement of the model and to explore the impact of the bilayer membrane in GnRHR conformation.

**Keywords:** Gonadotropin releasing hormone (GnRH); homology model; GnRH receptor; molecular dynamics; GPCR

## I. Introduction

The reproductive function and sexual maturation of human cells is regulated by hormones such as Gonadotropin-Releasing Hormone (GnRH) or Luteinizing–hormone–releasing hormone (LHRH).[1] GnRH is a decapeptide isolated from hypothalamic cells.[2] Contrary to other hormones, GnRH is released in pulses with frequencies varying from ~30 min to 120 min.[3] The hormone initiates a biochemical process by interacting with the Gonadotropin-Releasing Hormone Receptor (GnRHR) on the surface of cells. The activated biochemical pathway leads to the release of the Luteinizing Hormone (LH) and the Follicle–Stimulating Hormone (FSH).[1, 4, 5] Three isoforms of GnRH have been identified in vertebrates that contain conserved residues in the N– and C–terminal ends (Figure 1A).[6] GnRH isoforms I and II have been identified in most vertebrate species,[7] with GnRH I presenting the most diverse sequence.[8] The conserved residues of GnRH have been proposed to affect the binding of the peptide to GnRHR and the receptor's activation, while the variable residues in positions 5–8 have been linked to the function and ligand selectivity.[9-11] The suggested mode of action for GnRH involves distinct active conformations of GnRHR induced by the interactions with the GnRH isoforms. Furthermore, there is a very tight link between GnRH and GnRHR with the hormone regulating GnRHR expression via a negative feedback loop.[12]

**(A)**

***GnRH I*** (P)Glu$^{1}$-His$^{2}$-Trp$^{3}$-Ser$^{4}$-Tyr$^{5}$-Gly$^{6}$-Leu$^{7}$-Arg$^{8}$-Pro$^{9}$-Gly$^{10}$$NH_2$

***GnRH II*** (P)Glu$^{1}$-His$^{2}$-Trp$^{3}$-Ser$^{4}$-His$^{5}$-Gly$^{6}$-Trp$^{7}$-Tyr$^{8}$-Pro$^{9}$-Gly$^{10}$$NH_2$

***GnRH III*** (P)Glu$^{1}$-His$^{2}$-Trp$^{3}$-Ser$^{4}$-His$^{5}$-Asp$^{6}$-Trp$^{7}$-Lys$^{8}$-Pro$^{9}$-Gly$^{10}$$NH_2$

**(B)**

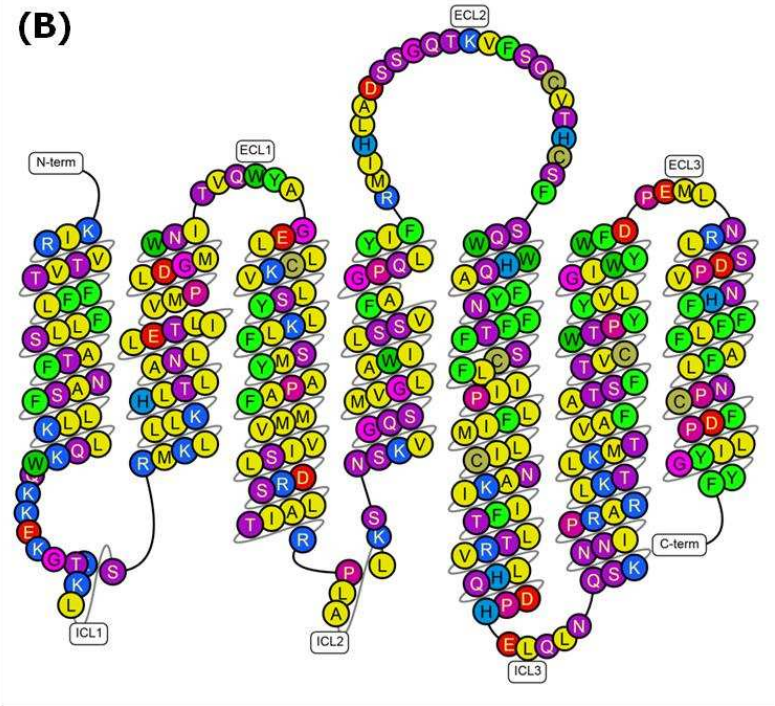



**Figure 1:** (A) The amino acid sequences of the three natural isoforms of GnRH, with the conserved residues coloured red and (B) schematic representation of GnRHR, showing the TM helical structures connected via extracellular (ECL) and intracellular (ICL) loops. Residues coloured green show increased binding/potency, red present reduced binding/potency and yellow have no effect (adapted from GPCR database: http://gpcrdb.org/).[13]

GnRH receptors are part of the rhodopsin family of G-Protein Coupled Receptors (GPCRs) and present in three subtypes (namely GnRHR I, II and III), similarly to GnRH. Studies have shown that the receptors from mammalian species present more than 80% residue similarity.[14-16] At the same time non–mammalian GnRHRs

present less similarity (~45%) with their mammalian counterparts.[17-19] The differentiation in the sequence, points to the divergence in the evolution of GnRHR in mammals. This observation is further supported by the loss of the C–terminal domain in GnRHR I as well as in the mammalian GnRHR II.[20] In human cells, GnRHR II expression has been silenced despite the conservation of GnRH II across species.[21] This phenomenon could be attributed to the possibility of GnRHR I to be activated by different GnRH isoforms. There are studies reporting that antagonists and GnRH II induce antiproliferative effects[22, 23] while GnRH III exhibits similar activity in apoptotic prostate cancer cells, both via GnRHR I activation.[22]

The rhodopsin family of GPCRs is composed of seven helical trans–membrane (TM) domains connected via extracellular (ECL) and intracellular (ICL) loops (Figure 1B). The TM domains of GnRHR have been implicated in conformational changes relevant for signal transduction, while the ECL domains have been involved in substrate binding.[24] Rhodopsin was the first GPCR structure solved by X–ray crystallography,[25] but several members of the family, including GnRH receptors, have no experimentally derived structure. Thus, any conformational and functional information for GnRHR is derived from modeling techniques and site–directed mutagenesis experimental data. The proposed theoretical models are based on both common residues and structural characteristics of GPCRs. The seven TM helical domains create a bundle that penetrates the cell membrane, adopting a similar conformation in most members of the rhodopsin family. The greatest variation is mostly observed in the ECL and ICL domains. In the rhodopsin family, residues $Asp^{2.50}$ (in TM2) and $Asn^{7.49}$ (in TM7) are conserved, but in GnRHR the positioning of these residues is inverted to $Asn^{2.50(87)}$ and $Asp^{7.49(319)}$.[26] This suggests a complementary role for these amino acids in the conformation and activation of receptor.[26]

Experimental data have suggest that residues on TM1 [$Arg^{1.35(38)}$], TM2 [$Asp^{2.61(98)}$/$Asn^{2.65(102)}$], TM3 [$Lys^{3.32(121)}$], TM5 [$Asn^{5.40(212)}$], TM6 [$Tyr^{6.58(290)}$] and TM7 [$Asp^{7.31(302)}$] are involved in the binding of the hairpin ($\beta$–turn) conformation of GnRH I.[27-29]

The close association of GnRH and GnRHR with various types of cancer highlights the importance of the receptor in new potential cancer treatments. Thus, the development of a valid homology model for the receptor may lead to a better understanding of its function. The application of theoretical approaches for the study of GnRHR and its interactions with GnRH will provide valuable information for the development of new and improved peptide and/or non-peptide mimetics of GnRH.

## II. Methods

### II.1 GnRH Receptor Homology model

The amino acid sequence of GnRH receptor I in FASTA format was employed as the search template for the construction of the model. The homology modeling process was performed with MODELLERv9.17 software,[30] as implemented in the PyMOL software package.[31] The crystal structures of human β2–adrenergic receptor (PDB code: 2rh1),[32] bovine (PDB code: 3dqb)[33] and squid (PDB code: 3ayn)[34] rhodopsins were implemented in the homology model process. Despite the low sequence similarity observed in the members of the rhodopsin family, it has been suggested that the seven TM helices share a similar orientation between the members of the family.[27, 28, 35] The structural similarities combined with conserved residues in the TM domains make possible the use of rhodopsin as a template for model building.[27, 28, 35] The residue sequences of the template proteins were aligned with the GnRHR I using the SALIGN function of MODELLER. Subsequently, the

homology mode was built based in the multiple sequence alignment result. The information regarding the presence of disulfide bonds ($Cys^{14}$–$Cys^{200}$ and $Cys^{114}$–$Cys^{196}$) have been incorporated in the construction of the model. Additionally, we took into account the presence of the potential salt bridge between residues $Lys^{3.32(121)}$ and $Glu^{2.53(90)}$.[36]

*Loop modeling and refining.* Most of the intracellular (ICL) and extracellular (ECL) loops of the receptor are small in size, 4-8 amino acids long (Figure 1B). Additionally, there are no information regarding their functionality in the binding of GnRH and the activation of the receptor. Thus, loops ICL1–3, ECL1 and ECL3 were not considered for loop modeling/refinement. On the other hand, the SuperLooper2 web server (http://proteinformatics.charite.de/ngl-tools/sl2/start.php)[37] has been implemented for the modeling of ECL2 (Figure 1B) which has a length of 22 amino acids. The pdb file of the constructed GnRHR I model was uploaded in the web server and the sequence of the loop was introduced in two segments, segment A ($Ile^{177}$-$Cys^{196}$) and segment B ($Cys^{196}$-$His^{207}$). This step was introduced in order not to change the position of $Cys^{196}$ that forms a disulfide bridge with $Cys^{114}$. Superlooper2 searches a database of more than 100,000 crystal structures and employs a geometric fingerprint matching method to evaluate the fitness of the modeled loop. The best candidate loop models are ranked based on sequence similarity, steric clashes and structural information. The models with the highest similarity and the lowest number of clashes were selected for ECL2 modeling.

**II.2 Model Refinement**

The homology models, with the refined ECL2 domain, were subjected to energy minimization with the AMBER14 software.[38] The parameters for GnRH receptor I were constructed with the ff14SB forcefield.[39] The total charge of the system was

neutralized by the addition of fifteen $Cl^-$ ions. The minimization process was performed in implicit solvent over 6 consecutive steps with the addition of restraints on the movement of atoms. The initial restraint constant was defined at 10.0 kcal $mol^{-1}$ $Å^{-2}$ and was subsequently reduced by 2 after each step. Every step of the minimization process involved 100 iterations (20 steps of steepest descent and 80 steps of conjugate gradient optimization).

**II.3 Molecular Dynamics Simulations**

The minimized homology model was subjected to all–atom molecular dynamic simulation. Since GnRH receptor is a trans–membrane protein the molecule was placed in a pre–equilibrated membrane to simulate its natural environment.

*Membrane building.* The pre–equilibrated membrane was built using the CHARMM–GUI membrane builder web tool (http://www.charmm-gui.org/?doc=input/membrane).[40] The particular web–tool allows the construction of the bilayer around the protein. The option of aligning the first principal axis along the Z axis was chosen for the positioning of the GnRHR. The bilayer composition of SDPC: SDPE: Cholesterol (a 2:2:1 mixture) was based on the work of Grossfield *et.al.* (2008)[41] on the molecule of rhodopsin. Since the AMBER14 lipid force field [42] does not support SD lipids, the composition of the bilayer was a 2:2:1 mixture of POPC: POPE: Cholesterol (total number of 125 residues per layer).

*MD simulation set up.* The ff14SB forcefield[39] was implemented for the construction of GnRHR parameters, while the lipid parameters where constructed with the lipid14 forcefield.[42] The parameters for the cholesterol molecule were obtained from Madej *et.al.* (2015).[43] The total charge of the system (GnRHR incorporated in the lipid membrane) was neutralized with the addition of 4 $K^+$ and 19 $Cl^-$ ions, since the partial charge of the lipid molecules had to be taken into account. The model was incorporated

in a rectangular box with dimensions 107.059x 106.295x 91.705 Å and the TIP3P water model[44] was used to solvate the system. The system was subsequently minimized in two steps; during the first step (5000 iterations of steepest descent followed by 20000 iterations of conjugate gradient) the lipids and the receptor were restrained with a force constant of 10 kcal mol–1 $Å^{-2}$. The following step involved the minimization of the whole system (5000 iterations of steepest descent followed by 45000 iterations of conjugate gradient) with a restraint of 10 kcal $mol^{-1}$ $Å^{-2}$ on GnRHR. Following minimization, the system was heated to 310 K over a total of 500 ps, with weak positional restraints (force constant 10.0 kcal $mol^{-1}$ $Å^{-2}$) placed on the GnRHR and the membrane lipids, in two steps: (a) the increase of the temperature from 0 to 100 K over 100 ps and (b) the increase of the temperature from 100 K to 310 K (on increments of 50 K) over 400 ps under constant volume conditions. The heating was followed by a density equilibration (constant pressure) phase for 5 ns with weak positional restraints (force constant 5.0 kcal $mol^{-1}$ $Å^{-2}$) placed on the GnRHR. Finally, an equilibration simulation run was performed for 100 ns under the NVT ensemble with the GnRH receptor held in position with a force constant of 5.0 kcal $mol^{-1}$ $Å^{-2}$. The restraints were removed from GnRHR during the MD simulation production run. Two independent MD simulation runs (seeds) of 350 and 125 ns, respectively for a total time of 475 ns were carried out for the GnRHR–membrane model.

*Trajectory and Clustering analysis*. All analyses (e.g. rmsd, clustering calculations and hydrogen bond interactions), were performed using the cpptraj module [45] of the AMBER14. The clustering analysis was based on the hierarchical approach [46] with the RMSD employed as the distance metric (cutoff 2.5 Å). All graphics were designed with UCSF Chimera 1.11.2.[47]

## III. Results and Discussion

### III.1 Homology model

The initial models for the receptor were based on the crystal structures of human β2–adrenergic receptor, bovine and squid rhodopsins. For comparison reasons four homology models were built, the three based on the individual proteins while the fourth model was constructed on a combination of structural features form the three proteins (Figure S1). The different models of the receptor were superimposed on the Cα, N and C backbone atoms in order to verify the positioning of conserved residues such as $Asn^{2.50(87)}$, $Asp^{2.61(98)}$, $Asn^{2.65(102)}$, $Lys^{3.32(121)}$ and $Asp^{7.49(319)}$ (Figure S1A). The models based on human β2–adrenergic receptor (Figure S1A, orange) and on the combined of all crystal structures (Figure S1A, blue) position the specific residues with orientation inside the 7TM helical bundle. Furthermore, the positioning of residues $Asn^{2.50(87)}$ and $Asp^{7.49(319)}$ is inside the helical bundle with the side chains pointing towards each other (Figure S1A). The particular orientation has been proposed to stabilize the conformation of GnRHR via the formation of a water–mediated bridge between the residues.[26, 48] In the models based on bovine and squid rhodopsin (Figure S1A, gold and green, respectively) the positioning of these residues is towards the outside of the helical bundle (Figure S1A, dotted lines). This suggests that the models based on the two rhodopsins (bovine and squid) may not incorporate important functional and conformational characteristics reported for members of this GPCR family. Another aspect observed in the homology model of GnRHR I based on the bovine rhodopsin is the conformation of TM7 helix (Figure 2B, gold). The software models TM7 helix with a kink at right angles (93.02°) between backbone atoms C [$Leu^{7.43(314)}$], Cα [$Asn^{7.44(315)}$] and Cα [$Pro^{7.45(316)}$] (Figure S1B) in contrast to the other models that represent helix

TM7 with no kink at its lower portion (Figure S1A). Thus, the model based on the 3dqb does not offer a good representation for the conformation of the particular helix. This mis-representation causes a shift in the position of $Asp^{7.49(319)}$ (Figure S1A) preventing the formation of the water bridge essential for the conformational stability of GnRHR I.[48]

Additionally, a comparison was performed between the proposed model (Figure S1A, blue) and homology models reported in the GPCR database (http://gpcrdb.org/, Figure 2A).[13] The homology models reported in the GPCR database are based on the crystal structures of the mouse active $\mu$–opioid receptor (PDB code: 5c1m),[49] the human active OX2–orexin receptor (PDB code: 4s0v)[50] and the chimeric serotonin receptor (PDB code: 4iaq).[51] The comparison of the different models of GnRHR I suggests a close agreement despite the different templates. Residues $Asn^{2.50(87)}$ and $Asp^{7.49(319)}$ present their side chains inside the helical bundle cavity (Figure 2A) with small variations in their orientations (Figure 2A). Proteins in the rhodopsin family do not share great sequence similarity but they share common structural characteristics. Thus, it is safe to assume that the positioning of the particular side chain residues may vary between the different models as well as between the different template crystal structures.

*ECL2 loop refining and modeling.* ECL2 is highly variable in the different members of the family and the crystal structure of rhodopsin cannot be implemented as a template.[25, 27] In rhodopsin, ECL2 adopts an anti–parallel β–hairpin conformation which is not conserved in other members of the family. ECL2 flanks the trans–membrane domains TM4 and TM5 (Figure 2B) and has been implicated in ligand binding and receptor activation.[52-55] It has been proposed that ECL2 acts as a "lid" on top of the binding pocket of GPCRs and thus regulates ligand entry into the

receptor.[53, 54] Despite the presence of the disulfide bridge between residues $Cys^{114}$, located on top of the TM3 domain, and $Cys^{196}$ (Figure 2B) the loop retains its freedom of movement. The disulfide bond between the two Cys residues has also been proposed to drive the proper folding and activation of GPCRs.[54, 55] The only feature conserved in most members of the rhodopsin family is the Cys–Cys bond which positions the loop on top of the TM–bundle further supporting the role of ECL2 as a regulator for ligand entry and binding.[27] Employing this feature in the loop modeling (see Methods section II.1) the 177-196 segment (Figure 2B, blue) was modeled on the human Angtiotensin receptor (PDB code: 4zud) and the 197-207 segment (Figure 2B, magenta) was modeled on a segment from cytochrome C (PDB code: 3cxh). The sequence similarity with the respective GnRHR segments was approximately 45% for both templates. Despite not having the highest similarity index compared to other templates, the template structures presented the lowest number of steric clashes with the rest of the residues of GnRH receptor and thus were selected as templates for the model of the ECL2 domain.

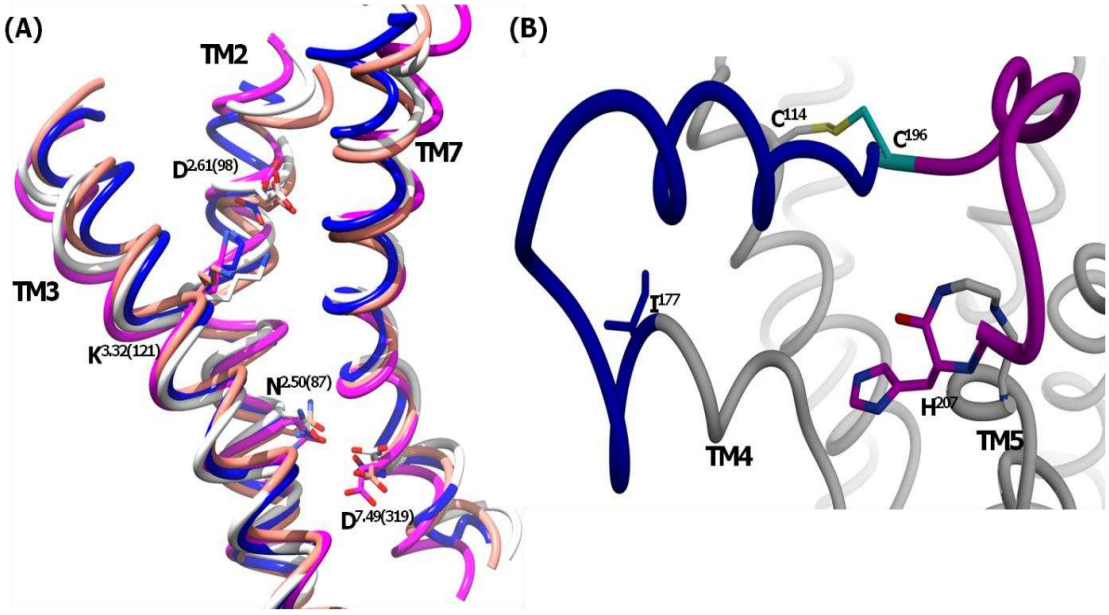


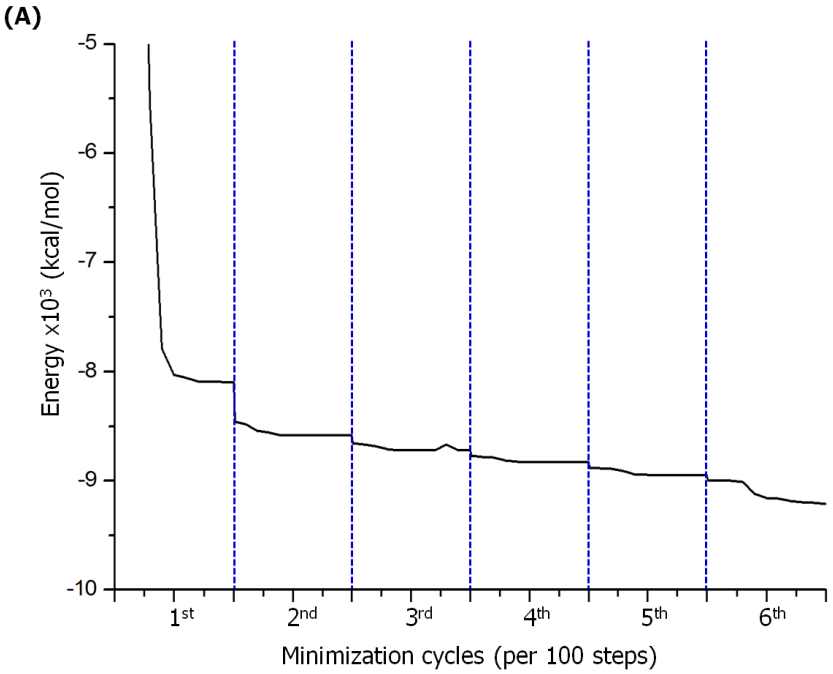

**Figure 2:** (A) Comparison of homology models for GnRHR I TM2, TM3 and TM7 helical domains, showing the side chain orientation for residues $Asn^{2.50(87)}$, $Asp^{2.61(98)}$,

$Asn^{2.65(102)}$, $Lys^{3.32(121)}$ and $Asp^{7.49(319)}$ represented as sticks. The color scheme of the different models is: blue for the model on all three crystal structures, pink for 5c1m, light grey for 4s0v, and orange for 4iaq and (B) ECL2 modeled conformation showing the anchor residues ($Ile^{177}$ and $His^{207}$) as sticks and the disulfide bridge between $Cys^{114}$ (in TM3 domain) and $Cys^{196}$.

**III.2 Model Refinement**

The development of a homology model, based on three different members of the rhodopsin family (described in Section II.2), poses certain challenges. Most members of the rhodopsin family of receptors share common structural features but lack great residue homology.[26, 48, 53] Thus, the development of a model may lead to steric clashes observed between residues that may not have been modeled correctly. This feature may be more prominent in the ECL and ICL domains since they show the greatest variability.[26] The best way to alleviate this potential problem is the refinement of the model with consecutive cycles of minimization employing positional restraints that are gradually abolished. In Figure 3A and Table S1 we present the changes in the potential energy ($E_{pot}$) of the proposed model, of GnRHR alone, over six cycles of minimization. The small number of steps (100) per cycle along with the imposition of restraints –that were gradually abolished (Table S1)– was employed in order for the minimized structure to retain the basic features of the model, while resolving any potential steric clashes. The graph (Figure 3A) presents the decrease in $E_{pot}$ over each cycle, with a very steep drop in the potential energy at the beginning of the first cycle and a further sharp drop at the start of the second cycle (Figure 3, Table S1). The largest deviation during the refinement process is observed on the ECL2 domain (Figure 3B). The RMSD between the ECL2 domain at the beginning of the

minimization process (Figure 3B, white) and at the end (Figure 3B, dark purple) is 0.5 Å compared to the RMSD of the whole GnRHR (0.2 Å) and that of the other ECL and ICL domains. The greatest difference in ECL2 is mainly focused on the positioning of residues Val$^{192}$, Ser$^{194}$, His$^{199}$, Trp$^{206}$, His$^{207}$ and Phe$^{210}$ (Figure 3B). The side chains of these residue deviate from their original positions and this change is most prominent of the conformation of the rings for His$^{199}$, Trp$^{206}$, His$^{207}$ and Phe$^{210}$.



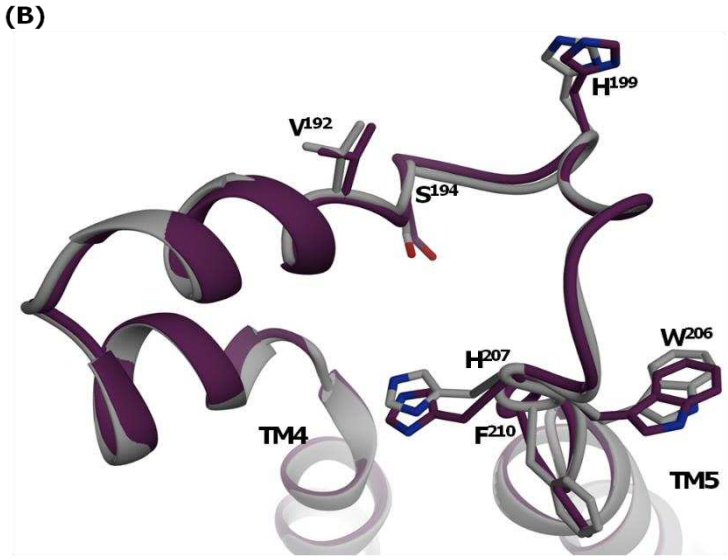



**Figure 3: (A)** Graph of $E_{pot}$ for the GnRHR homology model after six sequential minimization cycles and **(B)** the comparison of the ECL2 domain at the beginning (white) and the end (dark purple) of the refinement process.

### III.3 Molecular Dynamics

The construction of the homology model provides us with information regarding common structural features of members of the same family of proteins. Despite these similarities, the localization and function of these proteins may affect their conformation. Thus, MD simulations offer the tools to analyze the effect of the environment. In our case the homology model embedded in a bilayer, was subjected to two independent MD simulation runs of 350 and 125 ns, respectively (total time: 475 ns). Since GnRHR is present in various different cell types, there is no extensive information on its membrane environment. In order to explore the effect of the membrane in GnRHR conformation, we employed a membrane model developed for rhodopsin MD simulations.[41] The choice of the particular membrane model was based on the homology similarity of rhodopsin with GnRHR (both members of class A GPCRs) and the lipid parameters employed in the AMBER14 force field (see also section II.3 in Methods). The analysis of the RMSD values, for the Cα atoms (Figure S2A) shows that the movement of the Cα atoms of GnRHR deviates rapidly during the first few 5 ns of the simulation time compared to the reference structure (homology model). After this initial deviation, the Cα atoms of the receptor do not present great fluctuations during the rest of the simulation (Figure S2) and the RMS values are centered on 5 Å. This small deviation in the receptor movement can be attributed to the close packing imposed by the bilayer. Since the majority of GnRHR residues are found in the TM domains, their movement would be restricted by the close proximity of the

lipids in the bilayer. The changes in the atomic positional fluctuations of the residues are depicted in Figure 4, where the residues located in the trans-membrane helices present smaller deviation from their original conformation compared to the ECL and ICL domains.

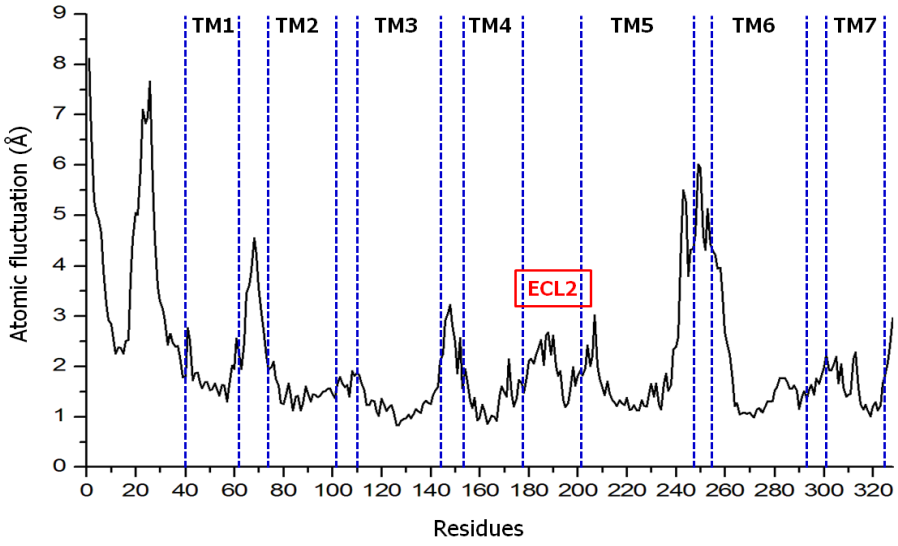



**Figure 4:** Positional atomic fluctuations (in Å) for the backbone atoms (N, C, O and Cα) of the GnRHR residues compared to the initial homology model over the MD simulation time. The dotted lines define the positions of the residues for the TM domains of the receptor.

Furthermore, the RMSD values for the backbone atoms of the TM domains were calculated in reference to the initial conformation of the homology model (Figure 6, Table S2). The results show small changes in the backbone conformation for TM domains 1–4 (Figure 5). The residues in these helices do not deviate extensively from their initial conformation with the maximum RMSD values reaching 2.96 Å for TM1 (Table S2) and the mean values for the four helices being 1.97 Å (for TM1), 1.34 Å (for TM2), 1.61 Å (for TM3) and 2.61 Å (for TM4) (see also Table S2). This small change

is mirrored in the positional fluctuations of the residues (Figure 4), where the mean atomic fluctuation for the residues in TM1 through TM4 is less than 2.0 Å (Table S2). The largest conformational changes during the MD simulation are focused on the residues of TM5 (Figure 5). For the first 50 ns there are extensive conformational changes in the TM5 reaching values of up to 5 Å before the changes reach equilibrium around 4.5 Å (Figure 5). Despite a small jump in the RMS between 180 and 200 ns, no major conformational changes are reported during the simulation time. These changes could be attributed to the large number of residues comprising the TM5 helix, as well as to the anchor segments of the domain (Figures 1 and 4). TM5 helix is located between the ECL2 and ICL3 domains. Both these domains are more mobile and thus the TM5 helix residues anchored to both ECL2 and ICL3 show higher positional fluctuations (Figure 4) with a mean atomic fluctuation of 2.01 Å. This observation may also account for the greater conformational changes observed for the TM6 helix (Figure 5). The anchoring segment of TM6 with ICL3 –which is highly mobile– (Figures 1 and 4), presents higher positional fluctuations that may affect the conformations of the helical domain leading to the increased RMSD values (mean RMSD 3.5 Å, Table S2). Finally, the TM7 domain shows smaller conformational changes over time (Figure 5) compared to domains TM5 and TM6. The average RMSD value is 2.90 Å (Table S2) with the maximum value of 4.00 Å being reached at 90 ns (Figure 5) before dropping and stabilizing at the average of 3.85 Å until the end of the simulation. This small conformational change is also mirrored on the positional fluctuations (Figure 4) with the mean for all the residues of the TM7 helix being 1.63 Å (Table S2). As expected, the atomic fluctuations of the ICL and ECL domains are higher than those observed for the TM domains (Figure 4) due to lack of any secondary structural features. Moreover, their positioning is outside the membrane bilayer and thus no restrictions, due to close

packing, are imposed on the residues. ICL1 and ICL3, along with the N-terminal domain, present the highest fluctuations despite their smaller size compared to ECL2 (Figure 4). The low fluctuations of the ECL2 domain could be attributed to the secondary structure elements observed (α–helix, Figure 3) and the presence of the disulfide bridge between $Cys^{114}$ and $Cys^{196}$.

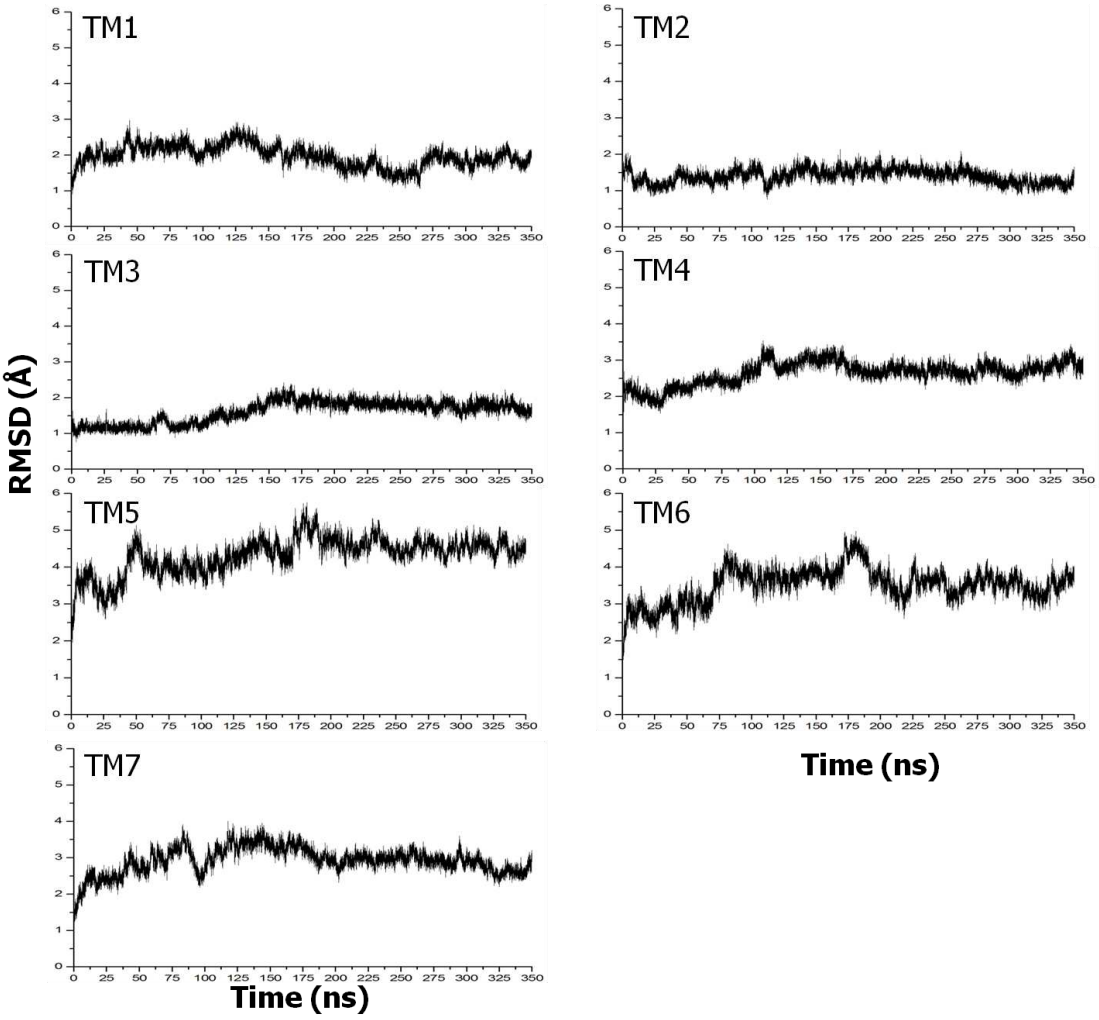



**Figure 5:** RMS deviations for the backbone atoms (N, C, O and Cα) of the residues comprising the seven TM domains compared to the initial conformation of the homology model over the MD simulation time. The differences of the RMS values over time present the conformational changes observed in the TM domains.

*Clustering analysis.* The clustering analysis performed on the MD simulations revealed the presence of five clusters for GnRHR (Table S3, Figure S2B). The dominant cluster 1 is present for 37% of the simulation time, cluster 2 is present for 24% of the time, while the other three clusters account for approximately the rest of the MD simulation time. The superimposition of the different clusters with the initial homology model shows the structural differences in the various TM domains observed over the course of the simulation (Figure 6). The comparison of the MD results reveals the small changes in domains TM1–4 (Figure 6A), while the greatest conformational changes are observed mostly on TM helices 5 and 6 (Figure 6B). As depicted in Figure 6B (grey, bottom), the TM5 and TM6 conformations deviate from the initial structure (Figure 6B, magenta). This result is in accordance with the RMS deviations observed for the two helices (Figure 5). Furthermore, it is important to note that the particular segments of the two helices are the anchor segments with the ICL3 domain connecting the two, which presents large atomic positional fluctuations (Figure 4).

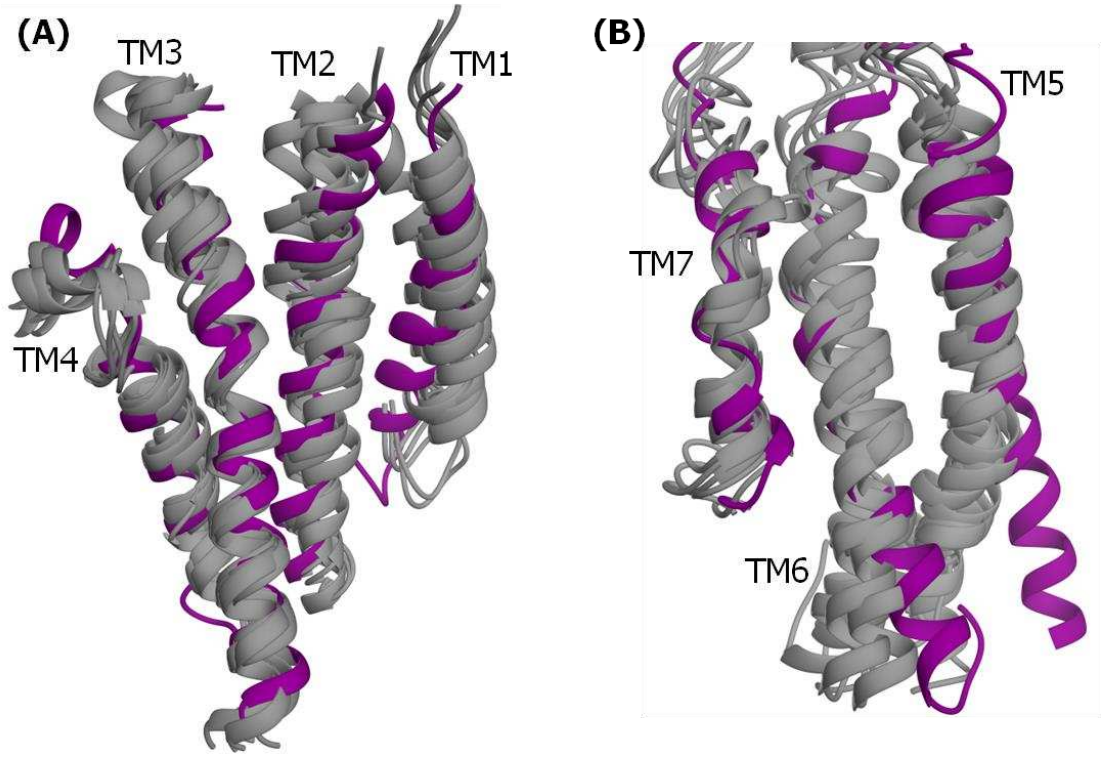


**Figure 6:** Superimposition of the different MD clusters (white) with the initial homology model (magenta): (A) TM domains 1 to 4 and (B) TM domains 5 to 7.

The clustering analysis revealed also the conformational changes observed for the amino acids important in GnRH binding and receptor activation –TM2 [$Asp^{2.61(98)}$/$Asn^{2.50(87)}$], TM3 [$Lys^{3.32(121)}$] and TM7 [$Asp^{7.49(319)}$]– (Figure 7). As expected, MD simulations create a dynamic picture of the molecule recording the different conformational changes over time. In this case, we report the changes in the distances between TM2 [$Asp^{2.61(98)}$]–TM3 [$Lys^{3.32(121)}$] and TM2 [$Asn^{2.50(87)}$]–TM7 [$Asp^{7.49(319)}$] (Figure 7). The distance between the Cα atoms of $Asp^{2.61(98)}$ and $Lys^{3.32(121)}$ is decreasing over the course of the MD production run (Figure 7A, right panel). This suggests a movement of the two helices towards each other in comparison to their initial position based on the homology model (distance of 10.3 Å). This movement could be readily explained by the effect of the lipid molecules that are in close proximity to the TM helices. The close packing imposed by the bilayer may lead to the movement of amino acid towards each other and affect not only the conformational aspects of the GnRHR. This is in accordance with the results of the clustering analysis that shows small conformational changes in TM2 helix between the homology model and the MD simulations (Figure 6A, magenta and grey, respectively). On the other hand, the distance between helices TM2 and TM7 remains relatively unchanged during the MD simulation compared to the homology model with the distances being 8.4 and 7.9 Å, respectively (Figure 7B, right panel). These observations suggest that helix TM2 may play an important role GnRH binding by its induced conformational change on the specific region (Figures 6A and 7A).

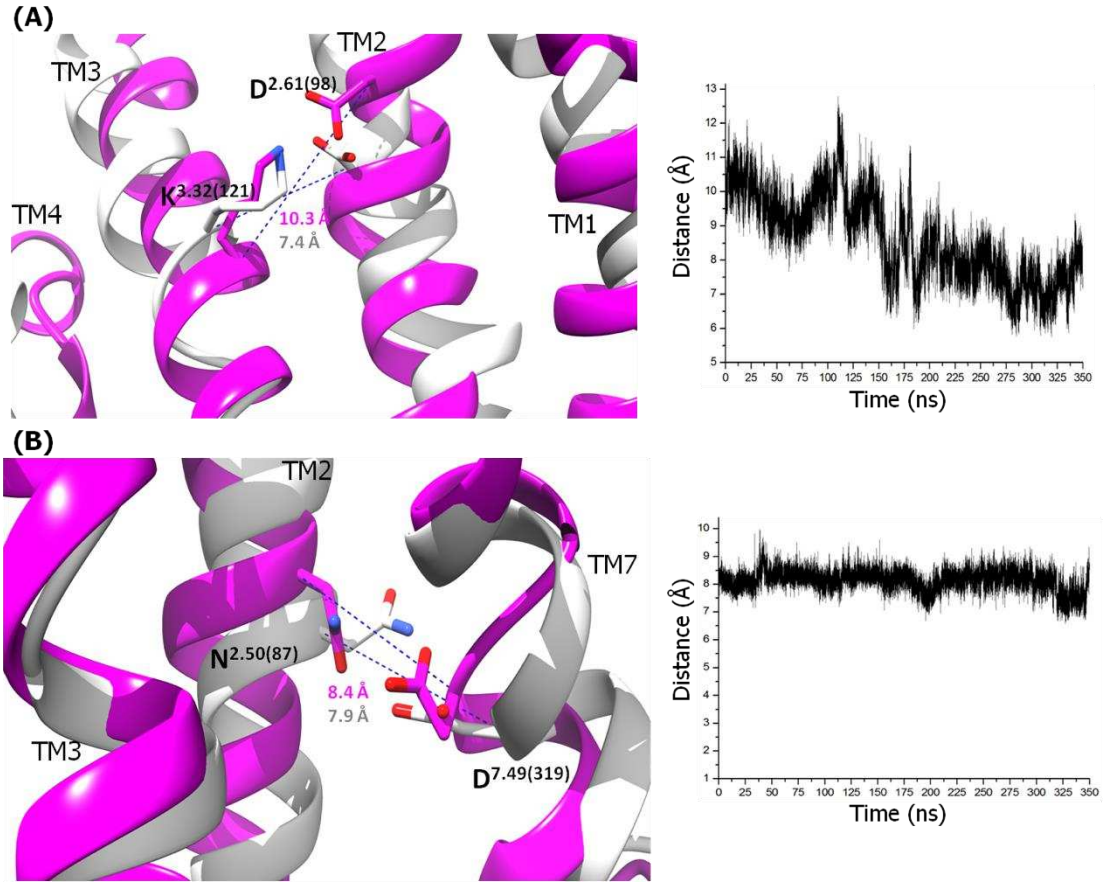



**Figure 7:** Distances between the Cα atoms of the dominant conformation (white) derived from the MD simulation with the initial homology model (magenta) for residues (A) $Asp^{2.61(98)}$ –$Lys^{3.32(121)}$ and (B) $Asn^{2.50(87)}$–$Asp^{7.49(319)}$. In the right panel we present the graphs depicting the changes of the respective distances over time (for the Cα atoms of the residues) during the MD simulation.

While the homology model provides the necessary structural information based on the different templates, it does not offer information regarding the effects of the membrane in its conformation. In this section we presented the dynamic features of conserved structural elements (see Figures 4, 5 and 7) of GnRHR. We record no major changes in the behavior of TM1–TM4 and TM7 (Figure 5), while a decrease in the distance between $Asp^{2.61(98)}$ –$Lys^{3.32(121)}$ is observed (Figure 7A). At the same time the orientation of residues $Lys^{3.32(121)}$ and $Glu^{2.53(90)}$ is not affected greatly during the MD

simulation (Figure S2B), a finding that is in agreement with other conformational studies of the receptor and other GPCRs.[36, 56]

### III.4 Conclusions

The homology model, reported in this study, was based on the conformational information derived from three different members of the rhodopsin family (namely the human β2–adrenergic receptor, the bovine and squid rhodopsins). The developed model shows great similarities in comparison with other models reported in the GPCR database (http://gpcrdb.org/, Figures 2 and S1). The structural similarities between our homology model and those already present in the GPCR database show that despite the differences in the residue sequence of the GPCRs, the conformational aspects required for the function of these receptors remain intact in the different members and affect the tertiary structure of the receptors. Additionally, MD simulations have been implemented in order to monitor the dynamic behavior of the molecule as it is embedded in a lipid bilayer. The analysis revealed that there are conformational changes in the domains of GnRHR (Figure 5) as well as in certain areas of the receptor that have been implicated in GnRH binding (Figure 7A). Most of these changes are expected since the close packing of the receptor inside the membrane may affect the movement of the different domains. But, the clustering analysis (Figure 6) showed that there are no great conformational changes in the receptor compared to the initial structure of the homology model. Even the orientation of residues does not change extensively and remains close to their initial position. The results reported in this paper corroborate the findings of different studies on the conformational aspects of class A GPCRs. Moreover, the comparison between our model and other homology models reported in the GPCR database (http://gpcrdb.org/, Figure 2A), [13] show that the main structural features of GnRHR are conserved despite the different templates employed.

This type of information can prove valuable in the understanding of GnRHR function and binding mode. Additionally, the homology model, along with the MD simulations, will provide the necessary template and stereochemical information for the design of new GnRHR potential agonist and antagonist peptide and non-peptide mimetic analogues.

**Acknowledgments**

We would like to acknowledge the funding by the States Scholarships Foundation (IKY) through the "Post-doctoral Financial Support action" part of the "Development of Human Resources, Education and Life-long learning action" co–financed by the European Social Fund (http://ec.europa.eu/esf/) and the Hellenic Republic.

## References

[1] Meysing, A.U., Kanasaki, H., Bedecarrats, G.Y., Acierno, J.S., Jr., Conn, P.M., Martin, K.A., et al. GNRHR mutations in a woman with idiopathic hypogonadotropic hypogonadism highlight the differential sensitivity of luteinizing hormone and follicle-stimulating hormone to gonadotropin-releasing hormone. The Journal of clinical endocrinology and metabolism. 2004, 89, 3189-98.

[2] Schally, A.V., Arimura, A., Kastin, A.J., Matsuo, H., Baba, Y., Redding, T.W., et al. Gonadotropin-releasing hormone: one polypeptide regulates secretion of luteinizing and follicle-stimulating hormones. Science. 1971, 173, 1036-8.

[3] Ferris, H.A., Shupnik, M.A. Mechanisms for pulsatile regulation of the gonadotropin subunit genes by GNRH1. Biology of reproduction. 2006, 74, 993-8.

[4] Pralong, F.P., Boepple, P.A., Conn, P.M., Whitcomb, R.W., Butler, J.P., Schoenfeld, D., et al. Contour of the GnRH pulse independently modulates gonadotropin secretion in the human male. Neuroendocrinology. 1996, 64, 247-56.

[5] Czieselsky, K., Prescott, M., Porteous, R., Campos, P., Clarkson, J., Steyn, F.J., et al. Pulse and Surge Profiles of Luteinizing Hormone Secretion in the Mouse. Endocrinology. 2016, 157, 4794-802.

[6] King, J.A., Hassan, M.F., Mehl, A.E., Millar, R.P. Gonadotropin-releasing hormone molecular forms in mammalian hypothalamus. Endocrinology. 1988, 122, 2742-52.

[7] Schneider, J.S., Rissman, E.F. Gonadotropin-releasing hormone II: a multi-purpose neuropeptide. Integrative and comparative biology. 2008, 48, 588-95.

[8] Perrett, R.M., McArdle, C.A. Molecular mechanisms of gonadotropin-releasing hormone signaling: integrating cyclic nucleotides into the network. Frontiers in endocrinology. 2013, 4, 180.

[9] Coccia, M.E., Comparetto, C., Bracco, G.L., Scarselli, G. GnRH antagonists. European journal of obstetrics, gynecology, and reproductive biology. 2004, 115 Suppl 1, S44-56.

[10] Salciccia, S., Gentilucci, A., Cattarino, S., Sciarra, A. GNRH-agonist or antagonist in the treatment of prostate cancer: a comparision based on oncological results. Urologia. 2016, 83, 173-8.

[11] Hannan, M.A., Kawate, N., Fukami, Y., Weerakoon, W.W., Bullesbach, E.E., Inaba, T., et al. Effects of long-acting GnRH antagonist, degarelix acetate, on plasma insulin-like peptide 3, testosterone and luteinizing hormone concentrations, and scrotal circumference in male goats. Theriogenology. 2017, 88, 228-35.

[12] Finch, A.R., Caunt, C.J., Armstrong, S.P., McArdle, C.A. Plasma membrane expression of gonadotropin-releasing hormone receptors: regulation by peptide and nonpeptide antagonists. Mol Endocrinol. 2010, 24, 423-35.

[13] Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsoe, K., Hauser, A.S., et al. GPCRdb: an information system for G protein-coupled receptors. Nucleic acids research. 2016, 44, D356-64.

[14] Perrin, M.H., Bilezikjian, L.M., Hoeger, C., Donaldson, C.J., Rivier, J., Haas, Y., et al. Molecular and functional characterization of GnRH receptors cloned from rat pituitary and a mouse pituitary tumor cell line. Biochemical and biophysical research communications. 1993, 191, 1139-44.

[15] Kakar, S.S., Musgrove, L.C., Devor, D.C., Sellers, J.C., Neill, J.D. Cloning, sequencing, and expression of human gonadotropin releasing hormone (GnRH) receptor. Biochemical and biophysical research communications. 1992, 189, 289-95.

[16] Weesner, G.D., Matteri, R.L. Rapid communication: nucleotide sequence of luteinizing hormone-releasing hormone (LHRH) receptor cDNA in the pig pituitary. Journal of animal science. 1994, 72, 1911.

[17] Sun, Y.M., Flanagan, C.A., Illing, N., Ott, T.R., Sellar, R., Fromme, B.J., et al. A chicken gonadotropin-releasing hormone receptor that confers agonist activity to mammalian antagonists. Identification of D-Lys(6) in the ligand and extracellular loop two of the receptor as determinants. The Journal of biological chemistry. 2001, 276, 7754-61.

[18] Seong, J.Y., Wang, L., Oh, D.Y., Yun, O., Maiti, K., Li, J.H., et al. Ala/Thr(201) in extracellular loop 2 and Leu/Phe(290) in transmembrane domain 6 of type 1 frog gonadotropin-releasing hormone receptor confer differential ligand sensitivity and signal transduction. Endocrinology. 2003, 144, 454-66.

[19] Alok, D., Hassin, S., Sampath Kumar, R., Trant, J.M., Yu, K., Zohar, Y. Characterization of a pituitary GnRH-receptor from a perciform fish, Morone saxatilis: functional expression in a fish cell line. Molecular and cellular endocrinology. 2000, 168, 65-75.

[20] Sefideh, F.A., Moon, M.J., Yun, S., Hong, S.I., Hwang, J.I., Seong, J.Y. Local duplication of gonadotropin-releasing hormone (GnRH) receptor before two rounds of whole genome duplication and origin of the mammalian GnRH receptor. PloS one. 2014, 9, e87901.

[21] Maggi, R., Cariboni, A.M., Marelli, M.M., Moretti, R.M., Andre, V., Marzagalli, M., et al. GnRH and GnRH receptors in the pathophysiology of the human female reproductive system. Human reproduction update. 2016, 22.

[22] Montagnani Marelli, M., Manea, M., Moretti, R.M., Marzagalli, M., Limonta, P. Oxime bond-linked daunorubicin-GnRH-III bioconjugates exert antitumor activity in

castration-resistant prostate cancer cells via the type I GnRH receptor. International journal of oncology. 2015, 46, 243-53.

[23] Wu, H.M., Huang, H.Y., Lee, C.L., Soong, Y.K., Leung, P.C., Wang, H.S. Gonadotropin-releasing hormone type II (GnRH-II) agonist regulates the motility of human decidual endometrial stromal cells: possible effect on embryo implantation and pregnancy. Biology of reproduction. 2015, 92, 98.

[24] Ferguson, S.S. Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. Pharmacological reviews. 2001, 53, 1-24.

[25] Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., et al. Crystal structure of rhodopsin: A G protein-coupled receptor. Science. 2000, 289, 739-45.

[26] Flanagan, C.A., Zhou, W., Chi, L., Yuen, T., Rodic, V., Robertson, D., et al. The functional microdomain in transmembrane helices 2 and 7 regulates expression, activation, and coupling pathways of the gonadotropin-releasing hormone receptor. The Journal of biological chemistry. 1999, 274, 28880-6.

[27] Forfar, R., Lu, Z.L. Role of the transmembrane domain 4/extracellular loop 2 junction of the human gonadotropin-releasing hormone receptor in ligand binding and receptor conformational selection. The Journal of biological chemistry. 2011, 286, 34617-26.

[28] Lu, Z.L., Coetsee, M., White, C.D., Millar, R.P. Structural determinants for ligand-receptor conformational selection in a peptide G protein-coupled receptor. The Journal of biological chemistry. 2007, 282, 17921-9.

[29] Flanagan, C.A., Rodic, V., Konvicka, K., Yuen, T., Chi, L., Rivier, J.E., et al. Multiple interactions of the Asp(2.61(98)) side chain of the gonadotropin-releasing

hormone receptor contribute differentially to ligand interaction. Biochemistry. 2000, 39, 8133-41.

[30] Webb, B., Sali, A. Comparative Protein Structure Modeling Using MODELLER. Current protocols in bioinformatics. 2014, 47, 5 6 1-32.

[31] Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.

[32] Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science. 2007, 318, 1258-65.

[33] Scheerer, P., Park, J.H., Hildebrand, P.W., Kim, Y.J., Krauss, N., Choe, H.W., et al. Crystal structure of opsin in its G-protein-interacting conformation. Nature. 2008, 455, 497-502.

[34] Murakami, M., Kouyama, T. Crystallographic analysis of the primary photochemical reaction of squid rhodopsin. Journal of molecular biology. 2011, 413, 615-27.

[35] Cvicek, V., Goddard, W.A., 3rd, Abrol, R. Structure-Based Sequence Alignment of the Transmembrane Domains of All Human GPCRs: Phylogenetic, Structural and Functional Implications. PLoS computational biology. 2016, 12, e1004805.

[36] Manilall, A., Stander, B.A., Madziva, M.T., Millar, R.P., Flanagan, C.A. Glu(2.53(90)) of the GnRH receptor is part of the conserved G protein-coupled receptor structure and does not form a salt-bridge with Lys(3.32(121)). Molecular and cellular endocrinology. 2019, 481, 53-61.

[37] Hildebrand, P.W., Goede, A., Bauer, R.A., Gruening, B., Ismer, J., Michalsky, E., et al. SuperLooper--a prediction server for the modeling of loops in globular and membrane proteins. Nucleic acids research. 2009, 37, W571-4.

[38] Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., et al. AMBER 14. University of California, San Francisco. 2014.

[39] Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of chemical theory and computation. 2015, 11, 3696-713.

[40] Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Davila-Contreras, E.M., et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. Journal of computational chemistry. 2014, 35, 1997-2004.

[41] Grossfield, A., Pitman, M.C., Feller, S.E., Soubias, O., Gawrisch, K. Internal hydration increases during activation of the G-protein-coupled receptor rhodopsin. Journal of molecular biology. 2008, 381, 478-86.

[42] Dickson, C.J., Madej, B.D., Skjevik, A.A., Betz, R.M., Teigen, K., Gould, I.R., et al. Lipid14: The Amber Lipid Force Field. Journal of chemical theory and computation. 2014, 10, 865-79.

[43] Madej, B.D., Gould, I.R., Walker, R.C. A Parameterization of Cholesterol for Mixed Lipid Bilayer Simulation within the Amber Lipid14 Force Field. The journal of physical chemistry. B. 2015, 119, 12424-35.

[44] Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983, 79, 926-35.

[45] Roe, D.R., Cheatham, T.E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. Journal of chemical theory and computation. 2013, 9, 3084-95.

[46] Shao, J., Tanner, S.W., Thompson, N., Cheatham, T.E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. Journal of chemical theory and computation. 2007, 3, 2312-34.

[47] Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al. UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004, 25, 1605-12.

[48] Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E.M., Shichida, Y. Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. Proceedings of the National Academy of Sciences of the United States of America. 2002, 99, 5982-7.

[49] Huang, W., Manglik, A., Venkatakrishnan, A.J., Laeremans, T., Feinberg, E.N., Sanborn, A.L., et al. Structural insights into micro-opioid receptor activation. Nature. 2015, 524, 315-21.

[50] Yin, J., Mobarec, J.C., Kolb, P., Rosenbaum, D.M. Crystal structure of the human OX2 orexin receptor bound to the insomnia drug suvorexant. Nature. 2015, 519, 247-50.

[51] Wang, C., Jiang, Y., Ma, J., Wu, H., Wacker, D., Katritch, V., et al. Structural basis for molecular recognition at serotonin receptors. Science. 2013, 340, 610-4.

[52] Conner, M., Hawtin, S.R., Simms, J., Wootten, D., Lawson, Z., Conner, A.C., et al. Systematic analysis of the entire second extracellular loop of the V(1a) vasopressin receptor: key residues, conserved throughout a G-protein-coupled receptor family, identified. The Journal of biological chemistry. 2007, 282, 17405-12.

[53] DeVree, B.T., Mahoney, J.P., Velez-Ruiz, G.A., Rasmussen, S.G., Kuszak, A.J., Edwald, E., et al. Allosteric coupling from G protein to the agonist-binding pocket in GPCRs. Nature. 2016, 535, 182-6.

[54] Avlani, V.A., Gregory, K.J., Morton, C.J., Parker, M.W., Sexton, P.M., Christopoulos, A. Critical role for the second extracellular loop in the binding of both orthosteric and allosteric G protein-coupled receptor ligands. The Journal of biological chemistry. 2007, 282, 25677-86.

[55] Goodwin, J.A., Hulme, E.C., Langmead, C.J., Tehan, B.G. Roof and floor of the muscarinic binding pocket: variations in the binding modes of orthosteric ligands. Molecular pharmacology. 2007, 72, 1484-96.

[56] Sakhteman, A., Khoddami, M., Negahdaripour, M., Mehdizadeh, A., Tatar, M., Ghasemi, Y. Exploring 3D structure of human gonadotropin hormone receptor at antagonist state using homology modeling, molecular dynamic simulation, and cross-docking studies. Journal of molecular modeling. 2016, 22, 225.